



TITLE:

虚探知事象を考慮した非線形搜索ゲームモデル (決定過程に関わる数理モデルの新たな展開と応用)

AUTHOR(S):

宝崎, 隆祐; 結家, 利之

CITATION:

宝崎, 隆祐 ...[et al]. 虚探知事象を考慮した非線形搜索ゲームモデル (決定過程に関わる数理モデルの新たな展開と応用). 数理解析研究所講究録 2013, 1857: 41-50

ISSUE DATE:

2013-10

URL:

<http://hdl.handle.net/2433/195261>

RIGHT:

虚探知事象を考慮した非線形搜索ゲームモデル

防衛大学校・情報工学科 宝崎 隆祐
Ryusuke Hohzaki
Department of Computer Science,
National Defense Academy
防衛省 結家 利之
Toshiyuki Kekka
Ministry of Defense

1 はじめに

搜索ゲームにおいて、搜索者は任意の位置へ手持ちの搜索資源を投入しつつ逃避者を探知しようとし、逃避者は搜索空間上における連続的な移動により搜索者からの逃避を図るゲームを特に搜索配分ゲーム (SAG) と呼ぶ [1, 3]. 搜索配分ゲームでは、搜索者の運動能力が相対的に高く、任意の場所へ移動して搜索可能であるという前提が重要であり、もし搜索者と逃避者の運動能力の間に明らかな差がなければ、彼らの戦略としてはともに移動戦略が重要になり、そのような搜索ゲームは逃避・搜索ゲーム (ESG) と呼ばれる. ESG に関しては、1 次元離散空間上で目標物と搜索者の双方の移動を議論したものに Washburn[12] があり、それは両プレイヤーが同一セルを選択することによる探知事象発生までの総搜索距離を支払いとする多段ゲームである. 潜伏する目標物に対して搜索コストを支払いとした研究としては Kikuta[7] がある.

さて、このような搜索ゲームに関する多くの研究では真目標のみの探知事象が扱われている. 搜索者の期待に反して真でない目標探知 (虚目標探知やノイズ発生) 事象を考慮したモデルが、虚探知モデルである. 虚探知発生は、真目標探知を目的とする搜索者の搜索には障害となる. 通常、虚探知は広域搜索により探知され、その後の精密調査 (精査) により真目標であるかそうでないかが判定される 2 段階搜索のプロセスを仮定する場合が多い. Stone ら [11] の研究は、虚目標も真目標と同じく搜索空間に分布し存在しており、投入された搜索努力に依存して探知されるものとして、真目標探知までの精査時間と搜索時間の和の期待値を最小にする搜索努力配分を論じたものである. その他にも、実体のないノイズ型虚探知を扱った Kisi[8] の研究や, Iida ら [6], Komiya ら [9] の研究がある.

以上のような虚探知モデルに関する研究は搜索者側の最適搜索計画のみを扱っていたのに対し、目標側の戦略も考慮した搜索ゲームを論じた初めての研究に Hohzaki[2, 4] がある. そこでは、虚探知発生の発生確率が投入された搜索資源に依存するとした複雑なモデルを考えているため、支払としての探知確率を厳密な式として取り扱わず、搜索資源戦略に対し線形な近似式を支払関数として採用している. 以上のように、その定式化の複雑さ故に、虚探知発生を陽に考慮した搜索ゲームは少なく、またプレイヤーの戦略に対し非線形となる支払関数を厳密に扱った研究は無い. 本研究では、虚探知発生を搜索の確率プロセスの中の事象として取り扱い、探知確率や搜索コストを考慮した利得を厳密な支払関数として採用して、SAG における搜索者の搜索資源配分戦略及び目標の移動戦略の最適解を導出する.

2 モデルの記述と事象のインスタンス

搜索者と目標が参加する 2 人ゼロ和ゲームで、虚探知事象の存在する次のような搜索配分ゲームを考える.

- A1. 搜索空間は、離散的なセル空間 $K = \{1, \dots, K\}$ と離散時点 $T = \{1, \dots, T\}$ から成る $K \times T$ である.

- A2. 目標は搜索空間上の1つのパスを選択することにより搜索者からの逃避を図る。離散搜索空間において考えられるパス全体の集合を Ω で表す。パス $\omega \in \Omega$ の時点 $t \in T$ での位置をセル $\omega(t)$ とする。
- A3. 搜索者はこの搜索空間上へ搜索努力を投入して目標探知を企図するが、搜索は時点 τ 以降にしか開始できない。その搜索可能な時間帯を $\hat{T} \equiv \{\tau, \dots, T\} \subseteq T$ で表す。非負である搜索努力量は、各時点 $t \in \hat{T}$ で総量 $\Phi(t)$ が利用可能である。時刻 t にセル i に投入する搜索努力量を $\varphi(i, t)$ で表す。ただし、努力投入には単位努力量当たりコスト $c_0(i, t) > 0$ を要する。時間 $t \in \hat{T}$ にセル $i \in K$ に目標が存在する場合、ここに投入される搜索努力量 $\varphi(i, t)$ による真目標の探知確率は次式で表される。

$$1 - \exp(-\alpha_i \varphi(i, t)) \quad (1)$$

α_i は、セル i における単位搜索努力量当たりの探知効率を示すパラメータである。

- A4. 搜索開始とともに真目標の探知、虚探知の2種類の探知事象が起こり得る。この探知事象のそれぞれが各時点において独立に高々1回生じ、時刻 t での虚探知の発生確率は Q_t とする。事象が発生しなければ、次の時点での搜索に移行する。探知事象が生じれば、その原因究明のため時点数 $t_f - 1$ を要する精密調査（精査）に入り、その間は搜索は実施できない。探知事象の真の原因は、精査時間経過後あるいは最終時点 T のどちらか早い時点で確実に診断される。それが真目標探知であればゲームは終了するし、虚探知のみの探知事象であれば、精査後再び搜索が再会される。
- A5. 搜索者は、真目標を探知した時点 t での価値 $V(t) \geq 0$ を正確な診断を得て獲得できる。ただし、この目標価値は時間とともに非増加的に変化するものとする。

$$V(t) \geq V(t+1) \quad (2)$$

- A6. ゲームは、真目標探知により、あるいは最終時点 T に達した時終了する。
- A7. ゲームの支払は搜索者の期待利得とする。利得とは真目標探知により得られる目標価値から、搜索に使用したそれまでの搜索コストを引いたもので定義する。搜索者はこれを大きくするように、目標は小さくするように行動する。

前提 A3 では、搜索者の搜索努力の配分戦略を $\varphi = \{\varphi(i, t), i \in K, t \in \hat{T}\}$ で表している。前提 A2 で定義したように、目標の純粋戦略は1つのパス ω の選択であるが、ここではパス ω をとる確率を $\pi(\omega)$ とする混合戦略 $\pi = \{\pi(\omega), \omega \in \Omega\}$ を考える。両戦略の実行可能領域は次式となる。

$$\Psi \equiv \left\{ \varphi \left| \sum_{i \in K} \varphi(i, t) \leq \Phi(t), t \in \hat{T}, \varphi(i, t) \geq 0, i \in K, t \in \hat{T} \right. \right\} \quad (3)$$

$$\Pi \equiv \left\{ \pi \left| \sum_{\omega \in \Omega} \pi(\omega) = 1, \pi(\omega) \geq 0, \omega \in \Omega \right. \right\} \quad (4)$$

両プレイヤーの戦略により、搜索開始後の各時点で、4種類の事象、状態が生じる。すなわち、搜索実施後の真目標探知が含まれる探知事象、搜索実施後の虚探知のみの探知事象、搜索実施後の無探知、及び精査状態の4種類である。単純化のため、それぞれを D, F, S, O で表すと、時間帯 \hat{T} の $T - \tau + 1$ 個の時点で4種類のいずれかの記号から成るインスタンスが確率的に生じることになる。ここでは真目標探知に興味があるため、その排反のインスタンス、すなわち、真目標探知 (D) のないインスタンスを羅列することを考える。前提 A4 から、生じ得るインスタンスとしては次が羅列できる。

- (1) 時点 τ には、S または F を置く。
- (2) 一般的な虚探知事象と精査の組合せは、F がありその後 $t_f - 1$ 回連続して O が来る。特殊なケースとして精査の途中で最終時点 T となり搜索が終了する場合があり、その場合は F が発生しその後連続して y 回（ただし、 $0 \leq y \leq t_f - 2$ ）O が並ぶ。
- (3) 以上2種類の記号の並び以外の時点では、S が置かれる。

最終時点を L とした検索時間帯 $[\tau, L]$ で、真目標探知の無いインスタンスの全集合 A_L は、次の2つの場合を考慮したアルゴリズムにより羅列できる。

まず、虚探知の後の精査の途中で最終時点 L に到達する場合であり、この最後の精査回数を y とすると、 y は $0 \leq y \leq t_f - 2$ の範囲内で値をとる。時点 τ からこの最後の虚探知が生起する直前の時点 $L - y - 1$ までの間に発生する虚探知回数 M は、各精査に続く精査回数 $t_f - 1$ を考慮して、高々 $\lfloor (L - y - \tau) / t_f \rfloor$ となる。したがって、 M は $0 \leq M \leq \lfloor (L - \tau - y) / t_f \rfloor$ の範囲で値をとる。虚探知とそれに続く精査のグループが M 回発生し、最後に $M + 1$ 回目の虚探知が時点数 y の精査を伴って最終時点まで続く。 $M + 1$ 回目の虚探知と精査以外には時点数 $L - \tau - Mt_f - y$ が存在するが、そこでは S が生じる。この S の1つのインスタンスは、 j 番目虚探知事象の前に存在する S の連続回数 x_j ($j = 1, \dots, M + 1$) を与えることによって作成できる。すなわち、

$$\sum_{j=1}^{M+1} x_j = L - \tau - Mt_f - y \quad (5)$$

を満たす任意の $\{x_j, j = 1, \dots, M + 1\}$ が1つのインスタンスを与え、その数は $L - \tau - M(t_f - 1) - y C_M$ 通りある。

次に、最終時点で精査が途中で中断しない場合であり、時点数 t_f を要する虚探知と精査から成る一連のグループが全時点数 $L - \tau + 1$ の中に収まる場合であり、虚探知の回数 M は $0 \leq M \leq \lfloor (L - \tau + 1) / t_f \rfloor$ の範囲で値をとる。残りの時点数 $L - \tau - Mt_f + 1$ における S の生起は、 j 番目の虚探知発生直前における S の連続発生回数 x_j ($j = 1, \dots, M$) と最後の S の発生回数 x_{M+1} を

$$\sum_{j=1}^{M+1} x_j = L - \tau - Mt_f + 1$$

となるように割り当てることで一意に決めることができ、その総数は $L - \tau - M(t_f - 1) + 1 C_M$ 通りとなる。この後者の場合は、実は前者の場合で $y = -1$ としたケースに等しい。以上の議論から、羅列 A_L のアルゴリズムは以下のとおりである。

- (i) y を $-1, 0, 1, \dots, t_f - 2$ とする。
- (ii) それぞれの y に対し、 M を $0, 1, \dots, \lfloor (L - y - \tau) / t_f \rfloor$ とする。
- (iii) y と M に対し、(5) を満たすベクトル $(x_j, j = 1, \dots, M + 1)$ を羅列する。

(i)~(iii) によって得られた $y, M, \{x_j\}$ の個々の組合せが1つのインスタンスを与える。以上により事象のインスタンスの作成法が分かったので、次節ではインスタンスの生起確率を求め、さらにゲームの支払である期待利得を導出する。

3 支払関数の導出

2節では、真目標探知の無い事象の羅列 A_L を作成した。そのインスタンスでは各時点 t での事象を記号 S, F, O で表していたが、ここでは支払関数を定式化するため、 $\{S, F, O\}$ の各記号をそれぞれ $\sigma(t) \in \{1, -1, 0\}$ とおいて、時点 $\tau, \tau + 1, \dots, T$ での全事象 A_T の個々のインスタンスを1つのベクトル $(\sigma(t), t \in \hat{T})$ で表現する。

時点 t での事象 $\sigma(t) \in \{1, -1, 0\}$ に対し、その生起確率として以下の一般式が得られる。

$$(1 - \sigma(t)Q_t - \delta_{\sigma(t), -1}) \exp(-|\sigma(t)|\alpha_{\omega(t)}\varphi(\omega(t), t))$$

ただし、 $\delta_{i,j}$ はクロネッカーのデルタであり、 $i = j$ のとき1を、 $i \neq j$ なら0をとる。この式を用いれば、時間帯 $[\tau, T]$ 間における真目標の探知確率 $P(\varphi, \omega)$ は、羅列 A_T の排反事象が生じる確率として次式で計算できる。

$$P(\varphi, \omega) = 1 - \sum_{\sigma \in A_T} \left[\prod_{t=\tau}^T (1 - \sigma(t)Q_t - \delta_{\sigma(t), -1}) \right] \exp\left(-\sum_{t=\tau}^T |\sigma(t)|\alpha_{\omega(t)}\varphi(\omega(t), t)\right) \quad (6)$$

また時間帯 $[\tau, t]$ 間において、インスタンス σ についての虚探知の発生・未発生に関連する確率 $Q_t(\sigma)$ と真目標に関連する探知確率 $P_t(\varphi, \omega, \sigma)$ 、及び搜索コスト $C_t(\varphi, \sigma)$ を求めると次式となる。

$$Q_t(\sigma) = \prod_{\zeta=\tau}^t \left(1 - \sigma(\zeta)Q_\zeta - \delta_{\sigma(\zeta), -1}\right) \quad (7)$$

$$P_t(\varphi, \omega, \sigma) = 1 - \exp \left(- \sum_{\zeta=\tau}^t |\sigma(\zeta)| \alpha_{\omega(\zeta)} \varphi(\omega(\zeta), \zeta) \right) \quad (8)$$

$$C_t(\varphi, \sigma) = \sum_{\zeta=\tau}^t |\sigma(\zeta)| \sum_{i \in K} c_0(i, \zeta) \varphi(i, \zeta) \quad (9)$$

時点 $t \in \hat{T}$ での探知により利得 $V(t) - C_t(\varphi, \sigma)$ が得られる場合と、最終時点まで探知がなく搜索コスト $C_T(\varphi, \sigma)$ のみを消費する場合とを考えると、期待利得は次式で与えられる。

$$\begin{aligned} R(\varphi, \omega) &= \sum_{t=\tau}^T \sum_{\sigma \in A_t} (V(t) - C_t(\varphi, \sigma)) Q_t(\sigma) (P_t(\varphi, \omega, \sigma) - P_{t-1}(\varphi, \omega, \sigma)) \\ &\quad - \sum_{\sigma \in A_T} C_T(\varphi, \sigma) Q_T(\sigma) (1 - P_T(\varphi, \omega, \sigma)) \end{aligned}$$

取り扱い易いように、この式に幾つかの変形を加え整理したものが次式である。

$$\begin{aligned} R(\varphi, \omega) &= \sum_{t=\tau}^{T-1} \sum_{\sigma \in A_{t+1}} \left\{ (V(t) - V(t+1)) + |\sigma(t+1)| \sum_{i \in K} c_0(i, t+1) \varphi(i, t+1) \right\} \\ &\quad \times Q_{t+1}(\sigma) P_t(\varphi, \omega, \sigma) + \sum_{\sigma \in A_T} V(T) Q_T(\sigma) P_T(\varphi, \omega, \sigma) - \sum_{\sigma \in A_T} C_T(\varphi, \sigma) Q_T(\sigma) \quad (10) \end{aligned}$$

式(2)による目標価値の非増加性及び(9)式による $C_t(\varphi, \sigma)$ の φ に関する線形性、更には(8)式で与えられる探知確率 $P_t(\varphi, \omega, \sigma)$ の φ に関する凹性から、利得 $R(\varphi, \omega)$ は変数 φ の関して狭義凹となる。最後に、目標の混合戦略 π により $R(\varphi, \omega)$ の期待値をとった期待支払は

$$\begin{aligned} R(\varphi, \pi) &= \sum_{\omega \in \Omega} \pi(\omega) R(\varphi, \omega) \\ &= \sum_{\omega \in \Omega} \pi(\omega) \left[\sum_{t=\tau}^{T-1} \sum_{\sigma \in A_{t+1}} \left\{ (V(t) - V(t+1)) + |\sigma(t+1)| \sum_{i \in K} c_0(i, t+1) \varphi(i, t+1) \right\} \right. \\ &\quad \times Q_{t+1}(\sigma) P_t(\varphi, \omega, \sigma) + \sum_{\sigma \in A_T} V(T) Q_T(\sigma) P_T(\varphi, \omega, \sigma) \left. \right] - \sum_{\sigma \in A_T} C_T(\varphi, \sigma) Q_T(\sigma) \quad (11) \end{aligned}$$

となるが、これは π に関して線形、 φ に関して狭義凹関数である。

4 期待支払最大化問題

ここでの問題は、目標戦略 π を所与として、(11)式による期待支払を最大化する搜索努力配分問題を解くアルゴリズムを提案することである。3節の議論から、最大化したい目的関数(11)式は φ に対して狭義凹関数であるから、次の期待利得最大化問題は唯一の最適解をもつ。

$$(P1) \max_{\varphi} R(\varphi, \pi) \quad s.t. \quad \sum_{i \in K} \varphi(i, t) \leq \Phi(t), \quad t \in \hat{T}, \quad \varphi(i, t) \geq 0, \quad i \in K, \quad t \in \hat{T}$$

問題 (P1) の最適解の必要十分条件は、ラグランジュ乗数 $\{\lambda(t), t \in \hat{T}\}$ と $\{\mu(i, t), i \in K, t \in \hat{T}\}$ を用いたラグランジュ関数

$$L(\varphi; \lambda, \mu) = R(\varphi, \pi) + \sum_{t \in \hat{T}} \lambda(t) \left(\Phi(t) - \sum_{i \in K} \varphi(i, t) \right) + \sum_{i \in K} \sum_{t \in \hat{T}} \mu(i, t) \varphi(i, t)$$

から導いた次の Karush-Kuhn-Tucker 条件 (KKT 条件) により与えられる.

$$\frac{\partial L}{\partial \varphi(i, t)} = \frac{\partial R}{\partial \varphi(i, t)} - \lambda(t) + \mu(i, t) = 0, \quad i \in K, \quad t \in \hat{T} \quad (12)$$

$$\lambda(t) \geq 0, \quad t \in \hat{T} \quad (13)$$

$$\mu(i, t) \geq 0, \quad i \in K, \quad t \in \hat{T} \quad (14)$$

$$\sum_{i \in K} \varphi(i, t) \leq \Phi(t), \quad t \in \hat{T} \quad (15)$$

$$\varphi(i, t) \geq 0, \quad i \in K, \quad t \in \hat{T} \quad (16)$$

$$\lambda(t) \left(\Phi(t) - \sum_{i \in K} \varphi(i, t) \right) = 0, \quad t \in \hat{T} \quad (17)$$

$$\mu(i, t) \varphi(i, t) = 0, \quad i \in K, \quad t \in \hat{T} \quad (18)$$

因みに $\partial R / \partial \varphi(i, t)$ を変数 $\varphi(i, t)$ を陽に外に出して整理すると、以下ようになる.

$$\frac{\partial R}{\partial \varphi(i, t)} = B(i, t) \exp(-\alpha_i \varphi(i, t)) + C(i, t) \quad (19)$$

ただし,

$$\begin{aligned} B(i, t) \equiv & \sum_{\omega \in \Omega_{it}} \pi(\omega) \left[\sum_{\zeta=t}^{T-1} \sum_{\sigma \in A_{\zeta+1}} \left\{ -\Delta V(\zeta) + |\sigma(\zeta+1)| \sum_{i \in K} c_0(i, \zeta+1) \varphi(i, \zeta+1) \right\} \right. \\ & \times |\sigma(t)| \alpha_i Q_{\zeta+1}(\sigma) \exp \left(- \sum_{\xi=\tau, \xi \neq t}^{\zeta} |\sigma(\xi)| \alpha_{\omega(\xi)} \varphi(\omega(\xi), \xi) \right) \\ & \left. + \sum_{\sigma \in A_T} V(T) |\sigma(t)| \alpha_i Q_T(\sigma) \exp \left(- \sum_{\xi=\tau, \xi \neq t}^T |\sigma(\xi)| \alpha_{\omega(\xi)} \varphi(\omega(\xi), \xi) \right) \right] \end{aligned} \quad (20)$$

$$C(i, t) \equiv \sum_{\sigma \in A_t} |\sigma(t)| c_0(i, t) Q_t(\sigma) \sum_{\omega \in \Omega} \pi(\omega) P_{t-1}(\varphi, \omega, \sigma) - \sum_{\sigma \in A_T} |\sigma(t)| c_0(i, t) Q_T(\sigma) \quad (21)$$

この式中で、 $\Omega_{it} \equiv \{\omega \in \Omega | \omega(t) = i\}$, $\Delta V(\zeta) \equiv V(\zeta+1) - V(\zeta)$ の記号を用いた. この KKT 条件を整理すれば、最適な搜索努力配分として次のような解析的な表現が得られる.

$$\varphi^*(i, t) = \frac{1}{\alpha_i} \left[\ln \frac{B(i, t)}{\lambda(t) - C(i, t)} \right]^+ \quad (22)$$

ただし、記号 $[x]^+ \equiv \max\{0, x\}$ を使った. また、この最適解による時点 t での搜索努力配分の総量

$$\sum_{i \in K | B(i, t) + C(i, t) > \lambda(t)} \frac{1}{\alpha_i} \left[\ln \frac{B(i, t)}{\lambda(t) - C(i, t)} \right]^+ \quad (23)$$

は、 $[0, \bar{\lambda}_t]$ (ただし、 $\bar{\lambda}_t \equiv \max_{i \in K} \{B(i, t) + C(i, t)\}$) の範囲の乗数 $\lambda(t)$ に対しては単調減少であり、それ以上の $\lambda(t) \geq \bar{\lambda}_t$ に対してはゼロとなる. さらに、 $B(i, t)$ 及び $C(i, t)$ には時点 t における搜索努力 $\{\varphi(i, t), i \in K\}$ は含まれないことをここで確認しておく.

以上の議論から、時点 t における最適解 $\{\varphi(i, t), i \in K\}$ を導出する次のような数値計算手順が構築できる。乗数 $\lambda(t)$ の最適値の探索を、 $\lambda(t) = 0$ から始める。これを (22) 式に用いて $\{\varphi(i, t), i \in K\}$ を計算する。もし $\sum_i \varphi(i, t) \leq \Phi(t)$ が成立すれば、最適解が求まったことになる。そうでなければ、 λ_t を上界として、 $\sum_{i \in K} \varphi(i, t) = \Phi(t)$ となる $\lambda(t)$ を求める。その際、(23) 式の単調減少性が利用できる。以上の手順を $t = \tau, \dots, T$ と変えながら実行してゆき、すべての $t \in \hat{T}$ に対し解の変更が無くなった時点で KKT 条件 (12)~(18) がすべて満たされたことになり、最適解 $\{\varphi^*(i, t)\}$ が導出されたことになる。各時点 t でこの計算手順を実施することにより期待利得 $R(\varphi, \pi)$ は単調に増加してゆくから、全体の数値計算アルゴリズムは最終的に最大期待利得に収束して終了する。以上をアルゴリズムとして書くと次のようになる。

最適搜索努力配分の導出アルゴリズム: $\Gamma(\pi)$

- (S1) 解 $\{\varphi(i, t), i \in K, t \in \hat{T}\}$ が収束するまで $t = \tau, \dots, T$ と変化させ、(S2)~(S3) を繰り返す。
 (S2) $\lambda(t) = 0$ とおき、(22) 式により $\{\varphi(i, t), i \in K\}$ を計算する。
 (S3) もし $\sum_{i \in K} \varphi(i, t) \leq \Phi(t)$ ならば終了し、(S1) に戻る。そうでなく $\sum_{i \in K} \varphi(i, t) > \Phi(t)$ ならば

次を実行することにより最適な乗数 $\lambda^*(t)$ を求め、(S1) に戻る。
 (i) $B(i, t) + C(i, t)$ の値をセル番号 $i \in K$ について小さい順に並べ、それを I_1, \dots, I_K とする。
 $\xi = 1$ とする。 $\lambda_0 = 0$ とおく。

(ii) $\lambda(t) = B(I_\xi, t) + C(I_\xi, t)$ を (22) 式に適用する。このことは次式により $\{\varphi(i, t), i \in K\}$ を計算することに等しい。

$$\varphi(i, t) = \begin{cases} 0, & i = 1, \dots, I_\xi, \\ (1/\alpha_i) \ln \{B(i, t) / (\lambda(t) - C(i, t))\}, & i = I_{\xi+1}, \dots, I_K \end{cases} \quad (24)$$

(iii) $\sum_{i \in K} \varphi(i, t) \leq \Phi(t)$ ならば、

$$\sum_{k=\xi+1}^K \frac{1}{\alpha_{I_k}} \ln \frac{B(I_k, t)}{\lambda^*(t) - C(I_k, t)} = \Phi(t)$$

を満たす最適な $\lambda^*(t)$ を、左辺の単調減少性を利用した区間 $[\lambda_0, \lambda(t)]$ での 2 分探索法により求める。その後、 $\lambda^*(t)$ を (24) 式に代入して最適解 $\{\varphi(i, t), i \in K\}$ を計算する。

そうでなければ、 $\lambda_0 = \lambda(t)$, $\xi = \xi + 1$ として (ii) に戻る。

5 均衡解の数値解法アルゴリズム

4 節では、目標戦略 π を所与として期待利得を最大化する搜索努力の最適配分を求める数値アルゴリズムを提案した。ここでは、(11) 式で与えられる期待支払をもつゲームを考え、搜索者及び目標の最適戦略とゲームの値を導出するやり方について議論する。

前述したように、期待支払は搜索者戦略 φ について狭義凹で、目標戦略 π については線形であるから、期待支払のミニマックス値とマックスミニ値は一致することはすでに知られている [10]。そこで、以下ではマックスミニ値を求めよう。マックスミニ最適化問題は次のように変形できる。

$$\max_{\varphi} \min_{\pi} R(\varphi, \pi) = \max_{\varphi} \min_{\pi} \sum_{\omega \in \Omega} \pi(\omega) R(\varphi, \omega) = \max_{\varphi} \min_{\omega \in \Omega} R(\varphi, \omega)$$

したがって上式は次の問題と同値であり、この凸計画問題を解けば、搜索者の最適戦略 φ^* が導出できる。

$$(P_M) \quad \max_{\varphi, \nu} \nu \quad (25)$$

$$s.t. \quad R(\varphi, \omega) \geq \nu, \quad \omega \in \Omega, \quad (26)$$

$$\sum_{i \in K} \varphi(i, t) \leq \Phi(t), \quad t \in \hat{T}, \quad (27)$$

$$\varphi(i, t) \geq 0, \quad i \in K, \quad t \in \hat{T}. \quad (28)$$

また問題 $\min_{\pi} \sum_{\omega} \pi(\omega) R(\varphi, \omega)$ では、探索者戦略 φ に対する π の最適反応は、 $\nu \equiv \min_{\omega} R(\varphi, \omega)$ に対し、

$$R(\varphi, \omega) > \nu \text{ ならば } \pi(\omega) = 0 \quad (29)$$

を満たすべきである。

探索者の最適戦略 φ^* に関しては4節の数値解法アルゴリズム $\Gamma(\pi)$ を、目標の最適戦略 π^* に関してはこの最適反応の条件を利用して、均衡解を導出するアルゴリズムを以下で提案する。ここでは、 π を変化させつつアルゴリズム $\Gamma(\pi)$ により最適反応戦略 φ_{π}^* を求め、最終的には π が φ に対する最適反応となるようにもってゆく。 π の制御には、 $\pi(\omega)$ を増加させると $R(\varphi_{\pi}^*, \omega)$ が大きくなる次の性質を用いる。

補題 1 (参考文献 [5]) ある $k \in \Omega$ について $\pi_1(k) = \pi(k) + \Delta\pi(k)$ とし、 k 以外の他のパス $k \neq \omega \in \Omega$ については $\pi_1(\omega) = \pi(\omega)$ として、 π を π_1 に修正してアルゴリズム $\Gamma(\pi_1)$ を適用した場合、得られる最適解 $\varphi_{\pi_1}^*$ によるパス k の支払 $R(\varphi_{\pi_1}^*, k)$ は、元の $R(\varphi_{\pi}^*, k)$ に比べ、 $\Delta\pi(k) > 0$ ならば増加し、 $\Delta\pi(k) < 0$ ならば減少する。

Proof: 次の関係が π と π_1 に対して常に成り立つ。

$$\begin{aligned} R(\varphi_{\pi}^*, \pi_1) &= \sum_{\omega \in \Omega} \pi(\omega) R(\varphi_{\pi}^*, \omega) + \Delta\pi(k) R(\varphi_{\pi}^*, k) \\ &= R(\varphi_{\pi}^*, \pi) + \Delta\pi(k) R(\varphi_{\pi}^*, k) \leq R(\varphi_{\pi_1}^*, \pi_1) = R(\varphi_{\pi_1}^*, \pi) + \Delta\pi(k) R(\varphi_{\pi_1}^*, k) \end{aligned}$$

故に、

$$0 \leq R(\varphi_{\pi}^*, \pi) - R(\varphi_{\pi_1}^*, \pi) \leq \Delta\pi(k) (R(\varphi_{\pi_1}^*, k) - R(\varphi_{\pi}^*, k))$$

となり、証明が終了する。 \square

補題 1 を用いて、 π を変化させながら均衡解を求める数値計算アルゴリズムが提案できる。

均衡解の導出アルゴリズム: Λ

- (E1) π を次により初期化する: $\pi(\omega) = 1/|\Omega|$. $l = 0$ とおく。
- (E2) 与えられた π に対し、アルゴリズム $\Gamma(\pi)$ から最適解 φ_{π}^* を計算する。
 π を $\sum_{\omega} \pi(\omega) = 1$ となるように正規化する。
- (E3) $\{R(\varphi_{\pi}^*, \omega), \omega \in \Omega\}$ の値を小さい順に、 $W_1 < W_2 < \dots < W_M$ のように並べる。ここで W_k は値 $R(\varphi_{\pi}^*, \omega)$ が k 番目に小さいものであり、その値をもつ ω の集合を $\Omega_k \equiv \{\omega \in \Omega | R(\varphi_{\pi}^*, \omega) = W_k\}$ とする。
 $\eta(\omega) = \pi(\omega)$ に対し $\pi \in \Pi$ と条件 (29) が成り立てば終了する。
- (E4) l が偶数ならある 1 つのパス $k \in \Omega_1$ に対し $\pi(k)$ を微少量 $\Delta\pi(k) > 0$ だけ増加させ、 l が奇数なら $k \in \arg \max_{\{\omega | \pi(\omega) > 0\}} R(\varphi_{\pi}^*, \omega)$ なるあるパス k に対し $\pi(k)$ に負の微少量 $\Delta\pi(k) < 0$ を加え、新しい π とする。
 $l = l + 1$ として、(E2) に戻る。

上記のアルゴリズム Λ では、最大の期待利得 $R(\varphi_{\pi}^*, \omega)$ をもつパス ω の選択確率 $\pi(\omega)$ を微減させることによりその値を小さくさせ、相対的に他のパスの期待利得を増大させる。また最小の期待利得をもつパスに対しては、その選択確率を微増させることで期待利得の増大と他のパスの期待利得の減少を行う。最終的に、条件式 (29) を満足する π^* と、それに対し最大期待利得を実現する探索者戦略 $\varphi_{\pi^*}^*$ が均衡解として得られ、アルゴリズムは終了する。

通常の数値計算アルゴリズムと同様に、許容誤差や $\Delta\pi(k)$ の大きさの設定によっては、アルゴリズムの収束の可否や速さが変わってくるが、その設定要領に関しては参考文献 [5] に詳しい。

表 1b. パス上の累積資源量と選択確率 (ケース 1)

Paths	1	2	3	4
累積資源量	5.0	5.719	5.045	4.589
選択確率	0.261	0	0.37	0.37

(2) 虚探知発生の影響 (ケース 2)

このケースは、ケース 1 に $Q_t = 0.5$ ($t \in \hat{T}$) 及び $t_f = 3$ の変更を加えたケースである。虚探知発生のため、ゲームの値はケース 1 の 5.8 から 3.3 となる。このケースの均衡解を表 2a と 2b に示した。虚探知発生事象は、それに続く精査により、計画した資源投入が実施されない可能性がある。このことは、交点への資源投入効果が低下することを示唆しており、表 1a に比較して表 2a では交点以外への資源投入が増えている。同時に、交点の多いパス 2 と 1 に対する目標の忌避度も低下し、それぞれの選択確率がケース 1 から増加し、特にパス 1, 3, 4 の選択確率は同じになる。事前の搜索資源配分計画で資源配分が予定されている・いないに関わらず、実際に起こるインスタンスで精査となればそこでは資源を使用することはできず、また資源投入コストも掛からないから、ケース 1 に比べ搜索資源の積極的な投入計画を取り易くなり、時点 $t = 1$ では $\Phi(t) = 1$ が全量投入される。

表 2a. 最適搜索資源配分 (ケース 2)

Cells \ t	1	2	3	4	5	6	7	8	9	10
1	0.05	0	0	0	0	0	0	0	0	0
2	0.414	1	0	0.842	0.586	0.549	0.772	0	0	0
3	0.536	0	1	0	0.218	0.366	0.047	1	1	1
4	0	0	0	0.158	0	0	0.181	0	0	0
5	0	0	0	0	0.196	0.085	0	0	0	0
投入総量	1	1	1	1	1	1	1	1	1	1

表 2b. パス上の累積資源量と選択確率 (ケース 2)

Paths	1	2	3	4
累積資源量	5.67	5.614	5.168	4.162
選択確率	0.327	0.019	0.327	0.327

7 おわりに

この研究では、従来からあまり取り扱われてこなかった虚探知事象を考慮した搜索配分ゲームを取り扱った。虚探知はどんなセンサーを用いた搜索活動にも必ずつきまとうといってよいが、その割にはこれを扱った研究の少ないのは、モデル化や解法が難しいからである。本論文では、虚探知事象を確率事象として精緻に網羅したモデル化や計算手法を用いて問題を 2 人ゼロ和ゲームに定式化し、それに非線形計画法を適用して均衡解を導いた。

虚探知のある搜索ゲームに関しては支払関数を線形式に近似した従来研究があるが、そこでは搜索資源量に虚探知発生確率が依存する関係も加味されている。本モデルでは厳密な非線形関数として支払を取り扱ったが故に、そのような依存性まではモデル化できなかった。また、虚探知事象の現実性という意味では、時点のみならずセルに依存して発生確率や精査に要する時点数が増えるモデルも重要である。以上の点が将来の課題である。

参考文献

- [1] A.Y. Garnaev, *Search Games and Other Applications of Game Theory*, Springer-Verlag, Tokyo, 2000.
- [2] R. Hohzaki, A Search Game with Several Types of False Contacts, *Nonlinear Analysis and Convex Analysis* (Edited by W. Takahashi and T. Tanaka), Yokohama Publishers, London, pp.59-79, 2004.
- [3] R. Hohzaki, Search allocation game, *European J. of Operational Research* **172** (2006), 101–119.
- [4] R. Hohzaki, Discrete Search Allocation Game with False Contacts, *Naval Research Logistics*, **54**(1), pp.46–58, 2007.
- [5] R. Hohzaki and K. Iida, A Solution for a Two-Person Zero-Sum Game with a Concave Payoff Function, *Nonlinear Analysis and Convex Analysis*, World Science Publishing Co., London, pp.157–166, 1999.
- [6] K. Iida, R. Hohzaki and K. Kaiho, Optimal Investigating Search Maximizing the Detection Probability, *J. of the Operations Research Society of Japan*, **40**(3), pp.294–309, 1997.
- [7] K. Kikuta, A Search Game with Traveling Cost, *J. of the Operations Research Society of Japan*, **34**(4), pp.365–382, 1991.
- [8] T. Kisi, Optimal Stopping of the Investigating Search, *Search Theory and Applications*(NATO Conference Series II-8), pp.255–260, Plenum Press, N.Y., 1979.
- [9] T. Komiya, K. Iida and R. Hohzaki, An Optimal Investigation in Two Stage Search with Recognition Errors, *J. of the Operations Research Society of Japan*, **49**(2), pp.130–143, 2006.
- [10] G. Owen, *Game Theory*, Academic Press, N.Y., 1995.
- [11] L.D. Stone, *Theory of Optimal Search*, pp.136–178, Academic Press, N.Y., 1975.
- [12] A.R. Washburn, Search-Evasion Game in a Fixed Region, *Operations Research*, **28**, pp.1290–1298, 1980.